

ĐẠI HỌC QUỐC GIA TP. HỒ CHÍ MINH
TRƯỜNG ĐẠI HỌC CÔNG NGHỆ THÔNG TIN
KHOA CÔNG NGHỆ PHẦN MỀM

Trần Đình Phương Linh - 22520778

Đặng Thị Ngọc Minh - 22520857

Nguyễn Khánh Huy - 22520566

Trần Bảo Phú - 22521104

CHUYÊN ĐỀ MOBILE AND PERVASIVE COMPUTING

SOULSPACE – ỨNG DỤNG HỖ TRỢ SỨC KHỎE

TINH THẦN CHO GEN Z

GV Hướng Dẫn: Ths. Nguyễn Tấn Toàn

TP. HỒ CHÍ MINH, 2025

MỤC LỤC

Chương 1. GIỚI THIỆU	1
1.1. Giới thiệu đề tài	1
1.2. Mục tiêu nhóm	1
1.3. Đối tượng người dùng	1
1.4. Chức năng trọng tâm	1
1.5. Tiêu chí ngoài phạm vi ứng dụng (Out-of-Scope)	2
Chương 2. GIẢI PHÁP KỸ THUẬT	3
2.1. Danh sách các Actor trong hệ thống	3
2.2. Usecase Diagram tổng thể	4
2.2.1. Danh sách các Usecase của User	5
UC-U01: Xác thực Tài khoản	5
UC-U02: Quản lý Hồ sơ cá nhân	5
UC-U03: Nhật ký Cảm xúc (Journal)	6
UC-U04: Cộng đồng Ấn danh	6
UC-U05: Bài test Tâm lý	7
UC-U06: Tư vấn Chuyên gia	7
UC-U07: Mental Tree	8
UC-U08: Nhắc nhở (Reminders)	8
2.2.2. Danh sách các Usecase của Expert	9
UC-E01: Đăng ký Chuyên gia	9
UC-E02: Quản lý Lịch làm việc	9
UC-E03: Quản lý Cuộc hẹn	10
UC-E04: Chat với Khách hàng	10
UC-E05: Viết bài PR Chuyên môn	11
UC-E06: Dashboard Thống kê	11
2.2.3. Danh sách các Usecase của Admin	12
UC-A01: Xác minh Chuyên gia	12
UC-A02: Kiểm duyệt Bài viết Expert	13
UC-A03: Quản lý Bài test Tâm lý	13
UC-A04: Kiểm duyệt Nội dung Cộng đồng	14
UC-A05: Xử lý Báo cáo (Report)	14
UC-A06: Dashboard Thống kê	15

2.3. Kiến trúc hệ thống	16
2.3.1. Kiến trúc tổng thể	16
2.3.2. Data Flow Diagram (DFD)	17
2.4. Activity Diagram	20
2.4.1. AD-01: Đăng ký Tài khoản User	20
2.4.2. AD-02: Tạo Nhật ký Cảm xúc (Journal)	21
2.4.3. AD-03: Đặt lịch Tư vấn Chuyên gia	22
2.4.4. AD-04: Quy trình phê duyệt Expert	23
2.4.5. AD-05: Làm Bài Test Tâm lý	24
2.4.6. AD-06: Quản lý Cuộc hẹn (Expert)	25
2.4.7. AD-07: Kiểm duyệt Nội dung Cộng đồng (Admin)	26
2.4.8. AD-08: Chat Realtime giữa User và Expert	27
2.5. Một số giải thuật liên quan	28
2.5.1. Thuật toán nhận diện cảm xúc sử dụng mô hình Deep Learning (Tích hợp vào tính năng Journal)	28
2.5.2. Thuật toán nhận diện văn bản Toxic (Tích hợp vào tính năng người dùng đăng bài)	29
2.5.3. Thuật toán liên quan đến hệ thống Cây Tinh Thần (User Tree System)	31
2.6. Danh sách các màn hình phác thảo dự kiến	34
2.6.1. Nhóm màn hình liên quan đến Xác thực (Authentication)	34
2.6.2. Nhóm màn hình liên quan đến màn hình chính	35
2.7. Công nghệ sử dụng	37
2.7.1. Backend (SoulSpace-Backend)	38
Core Framework & Runtime	38
Database & ORM	38
Authentication & Security	39
AI/ML Libraries	39
External Services Integration	39
Utilities	40
2.7.2. Mobile Apps (SoulSpace & SoulSpace-FE-Expert)	40
Core Framework	40
Navigation & Routing	41
Styling & UI	41
State Management & Data	41

Media & Device	42
Utilities	42
2.7.3. Admin Panel (SoulSpace-FE-Admin)	43
Core Framework	43
UI Components & Styling	43
Data Visualization	44
Form & Validation	44
Utilities	45
2.7.4. External Services & APIs	45
Chương 3. ĐÁNH GIÁ VÀ KẾT LUẬN	46
3.1. Đánh giá mức độ hoàn thành	46
3.2. Phương hướng trong tương lai	46
3.3. Kết luận	46

Chương 1. GIỚI THIỆU

1.1. Giới thiệu đề tài

SoulSpace là một ứng dụng mobile giúp người trẻ – đặc biệt là Gen Z – ghi lại cảm xúc, chăm sóc sức khỏe tinh thần, đồng thời kết nối với cộng đồng ẩn danh và chuyên gia tâm lý. Ứng dụng sử dụng AI kết hợp các bài test khoa học để phân tích trạng thái cảm xúc, đưa ra phản hồi nhẹ nhàng, định hướng người dùng từng bước tự chữa lành hoặc kết nối với hỗ trợ chuyên sâu nếu cần.

1.2. Mục tiêu nhóm

- Tạo ra một ứng dụng cung cấp không gian an toàn, riêng tư cho người trẻ được nói ra, được lắng nghe và chữa lành.
- Kết hợp giữa trí tuệ nhân tạo và liệu pháp tâm lý hiện đại, giúp mỗi người từng bước hiểu mình – yêu mình – chăm sóc mình.
- Ứng dụng không thay thế chuyên gia, nhưng là “người bạn đồng hành” đầu tiên và bền vững trong hành trình tinh thần.

1.3. Đối tượng người dùng

- **User:** Người dùng cần theo dõi cảm xúc và hỗ trợ tâm lý.
- **Expert:** Chuyên gia tâm lý đã qua xác minh, cung cấp dịch vụ tư vấn.
- **Admin:** Quản trị viên vận hành, kiểm duyệt nội dung và quản lý hệ thống.

1.4. Chức năng trọng tâm

- **Theo dõi & Phân tích:** Nhật ký cảm xúc tích hợp AI phân tích tâm trạng và Speech-to-Text.
- **Kết nối:** Cộng đồng chia sẻ ẩn danh và hệ thống đặt lịch/chat với chuyên gia.
- **Đánh giá & Duy trì:** Bài test tâm lý chuẩn hóa và Gamification (Mental Tree) giúp tạo thói quen tích cực.

1.5. Tiêu chí ngoài phạm vi ứng dụng (Out-of-Scope)

- Không hỗ trợ Video Call trực tiếp (chỉ Chat).
- Không tích hợp thanh toán thực tế (chỉ mô phỏng).
- Không thay thế các chẩn đoán y khoa chuyên sâu.

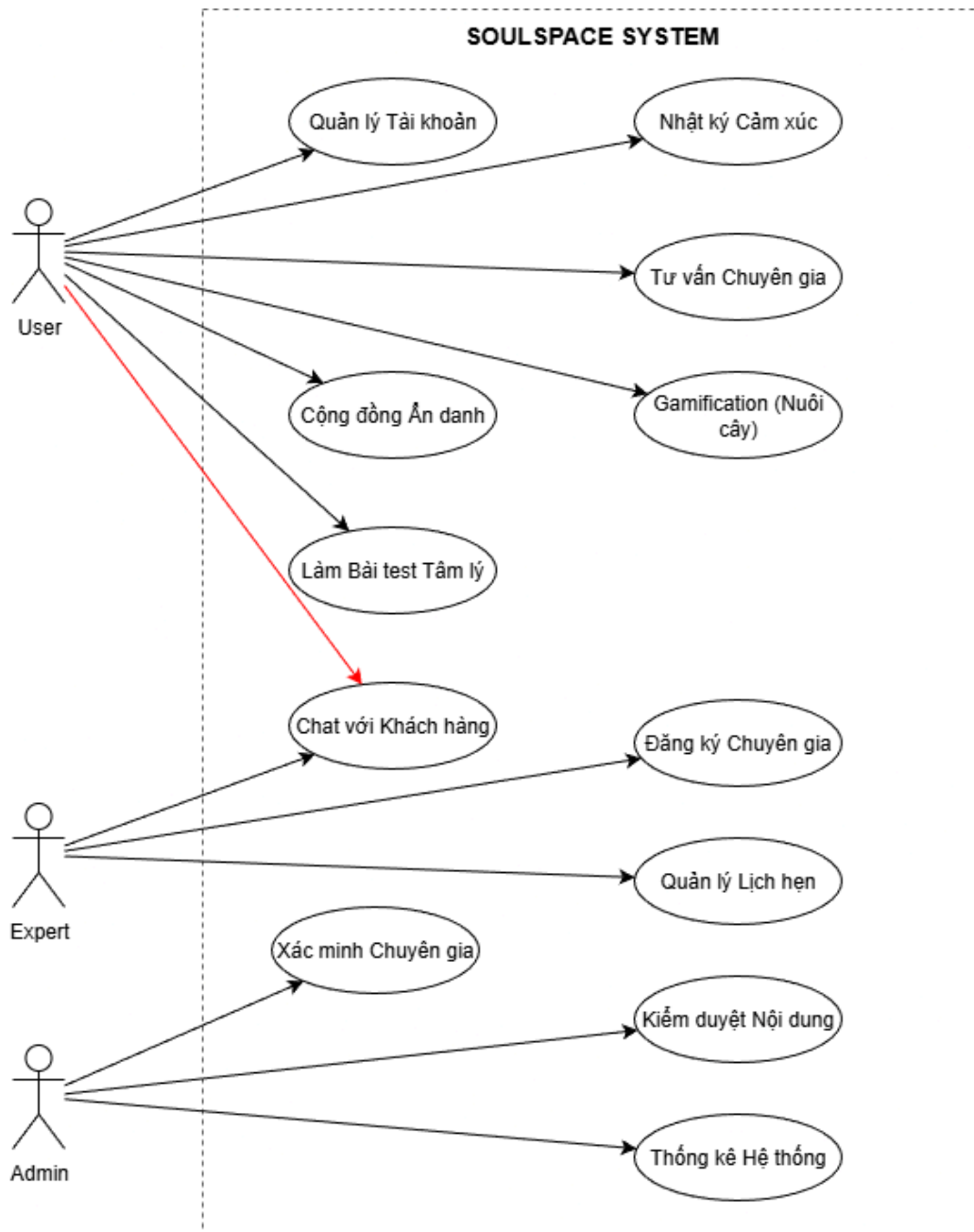
Chương 2. GIẢI PHÁP KỸ THUẬT

2.1. Danh sách các Actor trong hệ thống

Actor	Mô tả	Ứng dụng sử dụng
User	Người dùng cuối, sử dụng app để theo dõi sức khỏe tâm thần	SoulSpace Mobile App
Expert	Chuyên gia tâm lý, cung cấp dịch vụ tư vấn	SoulSpace Expert App
Admin	Quản trị viên hệ thống	SoulSpace Admin Panel

Bảng 2.1: Bảng danh sách các Actor trong hệ thống

2.2. Usecase Diagram tổng thể



Hình 2.1: Usecase Diagram tổng thể

2.2.1. Danh sách các Usecase của User

UC-U01: Xác thực Tài khoản

Thuộc tính	Mô tả
Actor	User
Mô tả	Người dùng thực hiện các thao tác xác thực để truy cập hệ thống
Tiền điều kiện	Ứng dụng đã được cài đặt
Luồng chính	1. User mở app → 2. Chọn Đăng nhập/Đăng ký → 3. Nhập thông tin → 4. Xác thực → 5. Truy cập hệ thống
Luồng thay thế	Quên mật khẩu: Nhập email → Nhận OTP → Xác thực → Đặt mật khẩu mới
Hậu điều kiện	User được xác thực và có thể sử dụng app

UC-U02: Quản lý Hồ sơ cá nhân

Thuộc tính	Mô tả
Actor	User (đã đăng nhập)
Mô tả	Quản lý thông tin cá nhân như avatar, tên người dùng
Tiền điều kiện	User đã đăng nhập
Luồng chính	1. Vào Settings → 2. Chọn Account → 3. Cập nhật thông tin → 4. Lưu

Thuộc tính	Mô tả
Hậu điều kiện	Thông tin được cập nhật trong hệ thống

UC-U03: Nhật ký Cảm xúc (Journal)

Thuộc tính	Mô tả
Actor	User (đã đăng nhập), AI System
Mô tả	Ghi lại cảm xúc hàng ngày với hỗ trợ AI phân tích
Tiền điều kiện	User đã đăng nhập
Luồng chính	1. Chọn tab Diary → 2. Chọn cảm xúc → 3. Thêm nội dung/ghi âm → 4. Thêm tag → 5. Lưu
Luồng mở rộng	AI tự động phân tích sentiment và phát hiện nội dung tiêu cực
Hậu điều kiện	Nhật ký được lưu, AI đã phân tích cảm xúc

UC-U04: Cộng đồng Ấn danh

Thuộc tính	Mô tả
Actor	User, AI System (kiểm duyệt)
Mô tả	Chia sẻ và tương tác trong cộng đồng hỗ trợ tâm lý
Tiền điều kiện	User đã đăng nhập
Luồng chính	1. Vào Community → 2. Xem/Tạo bài → 3. Like/Comment → 4. Chọn ấn danh

Thuộc tính	Mô tả
Luồng mở rộng	AI kiểm duyệt nội dung độc hại trước khi đăng
Hậu điều kiện	Bài đăng được tạo/tương tác thành công

UC-U05: Bài test Tâm lý

Thuộc tính	Mô tả
Actor	User
Mô tả	Thực hiện các bài test đánh giá sức khỏe tâm thần
Tiền điều kiện	User đã đăng nhập
Luồng chính	1. Vào Explore → 2. Chọn test → 3. Làm bài → 4. Xem kết quả
Luồng thay thế	Lưu tiến độ nếu chưa hoàn thành
Hậu điều kiện	Kết quả test được lưu và hiển thị

UC-U06: Tư vấn Chuyên gia

Thuộc tính	Mô tả
Actor	User, Expert
Mô tả	Đặt lịch và tư vấn với chuyên gia tâm lý
Tiền điều kiện	User đã đăng nhập, Expert đã được phê duyệt

Thuộc tính	Mô tả
Luồng chính	1. Tìm chuyên gia → 2. Xem lịch trống → 3. Đặt hẹn → 4. Thanh toán → 5. Chat
Hậu điều kiện	Cuộc hẹn được tạo, hai bên có thể chat

UC-U07: Mental Tree

Thuộc tính	Mô tả
Actor	User
Mô tả	Gamification - Chăm sóc cây ảo bằng các hành động tích cực
Tiền điều kiện	User đã đăng nhập
Luồng chính	1. Vào Plant → 2. Chọn hành động tích cực → 3. Tưới cây → 4. Nhận XP
Hậu điều kiện	Cây phát triển, XP tăng, streak được ghi nhận

UC-U08: Nhắc nhở (Reminders)

Thuộc tính	Mô tả
Actor	User
Mô tả	Tạo và quản lý các nhắc nhở chăm sóc bản thân
Tiền điều kiện	User đã đăng nhập
Luồng chính	1. Vào Reminders → 2. Tạo mới → 3. Đặt thời gian → 4. Lưu

Thuộc tính	Mô tả
Tùy chọn lặp	Một lần, Hàng ngày, Tùy chỉnh
Hậu điều kiện	Nhắc nhở được lên lịch

2.2.2. Danh sách các Usecase của Expert

UC-E01: Đăng ký Chuyên gia

Thuộc tính	Mô tả
Actor	Expert, Admin
Mô tả	Chuyên gia đăng ký tài khoản và chờ phê duyệt từ Admin
Tiền điều kiện	Expert có chứng chỉ hành nghề
Luồng chính	1. Đăng ký email/password → 2. Điền thông tin chuyên môn → 3. Upload avatar & chứng chỉ → 4. Chờ Admin phê duyệt
Thông tin yêu cầu	Họ tên, SĐT, Năm kinh nghiệm, Bio, Phòng khám, Địa chỉ, Giá tư vấn, Chứng chỉ
Hậu điều kiện	Tài khoản ở trạng thái "pending", chờ phê duyệt

UC-E02: Quản lý Lịch làm việc

Thuộc tính	Mô tả
Actor	Expert (đã phê duyệt)

Thuộc tính	Mô tả
Mô tả	Tạo và quản lý các slot thời gian có thể tư vấn
Tiền điều kiện	Expert đã được phê duyệt
Luồng chính	1. Vào Calendar → 2. Chọn ngày → 3. Tạo slot (giờ bắt đầu - kết thúc) → 4. Lưu
Hậu điều kiện	Slot hiển thị cho User khi đặt lịch

UC-E03: Quản lý Cuộc hẹn

Thuộc tính	Mô tả
Actor	Expert
Mô tả	Xử lý các yêu cầu đặt lịch từ User
Tiền điều kiện	Có User đặt lịch
Luồng chính	1. Xem yêu cầu pending → 2. Chấp nhận/Từ chối → 3. Quản lý upcoming
Tab hiển thị	Pending (chờ xử lý), Upcoming (sắp tới), Past (đã qua)
Hậu điều kiện	Cuộc hẹn được xác nhận hoặc từ chối

UC-E04: Chat với Khách hàng

Thuộc tính	Mô tả
Actor	Expert, User

Thuộc tính	Mô tả
Mô tả	Chat realtime với khách hàng đã đặt lịch
Tiền điều kiện	Có cuộc hẹn được xác nhận
Luồng chính	1. Mở danh sách chat → 2. Chọn conversation → 3. Gửi/Nhận tin nhắn
Tính năng	WebSocket realtime, typing indicator, online status
Hậu điều kiện	Tin nhắn được gửi và lưu trữ

UC-E05: Viết bài PR Chuyên môn

Thuộc tính	Mô tả
Actor	Expert, Admin
Mô tả	Expert viết bài chia sẻ kiến thức chuyên môn
Tiền điều kiện	Expert đã được phê duyệt
Luồng chính	1. Tạo bài mới → 2. Nhập tiêu đề, nội dung → 3. Upload ảnh → 4. Thêm hashtag → 5. Gửi duyệt
Hậu điều kiện	Bài viết ở trạng thái "pending", chờ Admin duyệt

UC-E06: Dashboard Thống kê

Thuộc tính	Mô tả
Actor	Expert

Thuộc tính	Mô tả
Mô tả	Xem tổng quan về hoạt động tư vấn
Tiền điều kiện	Expert đã được phê duyệt
Dữ liệu hiển thị	Số cuộc hẹn hôm nay, Số dư ví, Lịch hẹn upcoming
Hậu điều kiện	-

2.2.3. Danh sách các Usecase của Admin

UC-A01: Xác minh Chuyên gia

Thuộc tính	Mô tả
Actor	Admin
Mô tả	Xét duyệt hồ sơ đăng ký của chuyên gia
Tiền điều kiện	Có Expert đăng ký mới
Luồng chính	1. Xem danh sách pending → 2. Xem chi tiết hồ sơ → 3. Kiểm tra chứng chỉ → 4. Phê duyệt/Từ chối
Lọc theo	pending, approved, rejected
Hậu điều kiện	Expert được phê duyệt hoặc từ chối

UC-A02: Kiểm duyệt Bài viết Expert

Thuộc tính	Mô tả
Actor	Admin
Mô tả	Duyệt các bài PR của chuyên gia trước khi công khai
Tiền điều kiện	Expert đã submit bài viết
Luồng chính	1. Xem bài pending → 2. Đọc nội dung → 3. Phê duyệt/Từ chối
Hậu điều kiện	Bài viết hiển thị trong newsfeed (nếu approved)

UC-A03: Quản lý Bài test Tâm lý

Thuộc tính	Mô tả
Actor	Admin
Mô tả	CRUD các bài test đánh giá sức khỏe tâm thần
Tiền điều kiện	Admin đã đăng nhập
Luồng chính	1. Xem danh sách → 2. Tạo/Sửa/Xóa test → 3. Quản lý câu hỏi và đáp án
Thông tin test	Mã, Tên, Mô tả, Ngưỡng nghiêm trọng, Khuyến nghị, Hình ảnh, Câu hỏi
Hậu điều kiện	Test được cập nhật trong hệ thống

UC-A04: Kiểm duyệt Nội dung Cộng đồng

Thuộc tính	Mô tả
Actor	Admin, AI System
Mô tả	Giám sát và quản lý nội dung trong cộng đồng
Tiền điều kiện	-
Luồng chính	1. Xem tất cả posts → 2. Lọc theo status → 3. Xóa nội dung vi phạm
Trạng thái	Approved, Pending, Blocked
Hậu điều kiện	Nội dung được kiểm duyệt

UC-A05: Xử lý Báo cáo (Report)

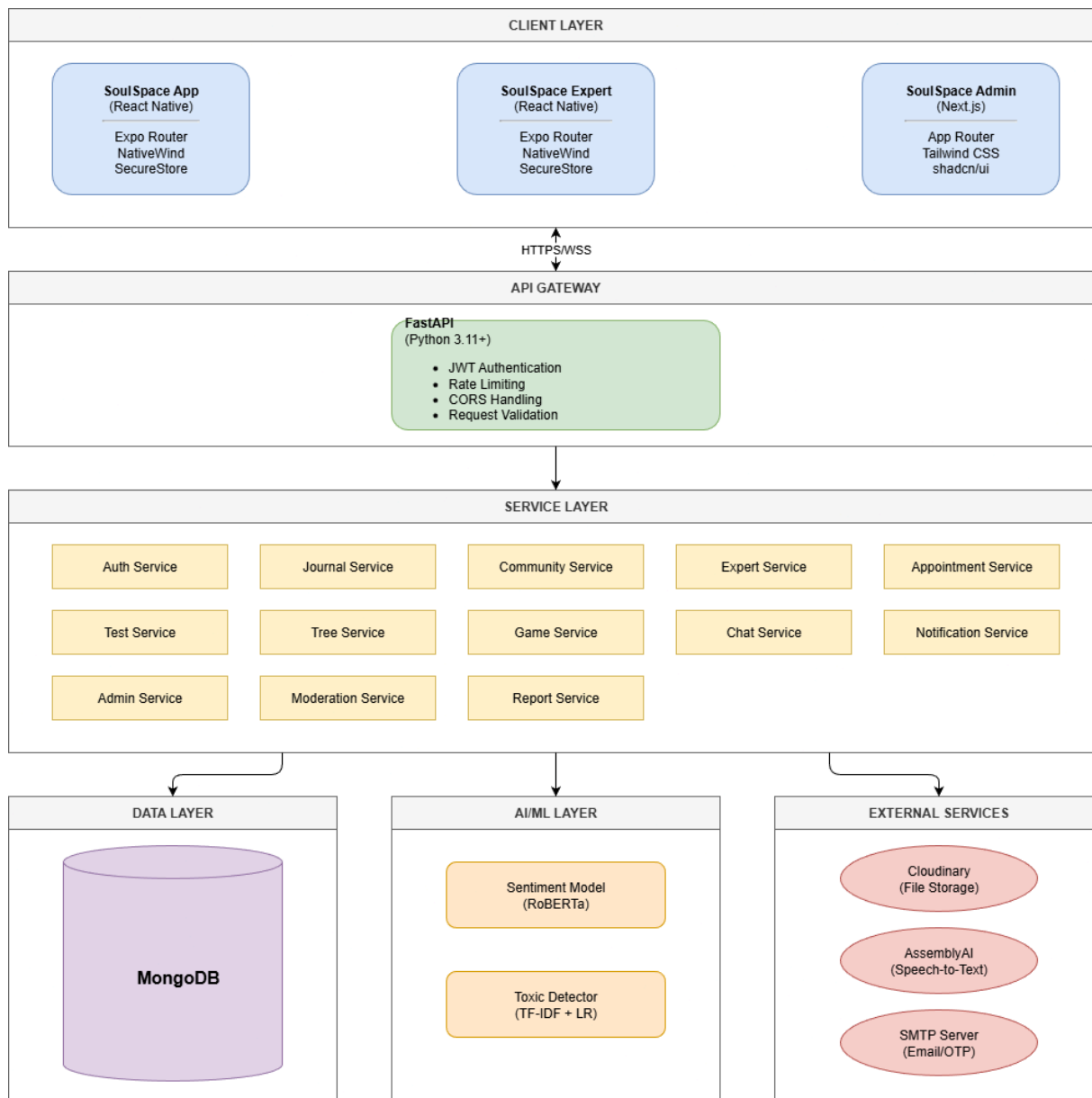
Thuộc tính	Mô tả
Actor	Admin
Mô tả	Xử lý các báo cáo vi phạm từ người dùng
Tiền điều kiện	User đã báo cáo nội dung
Luồng chính	1. Xem report → 2. Xem nội dung bị báo cáo → 3. Quyết định xử lý
Hành động	delete_content, warn_user, dismiss
Hậu điều kiện	Report được xử lý

UC-A06: Dashboard Thống kê

Thuộc tính	Mô tả
Actor	Admin
Mô tả	Xem tổng quan về hoạt động của hệ thống
Tiền điều kiện	Admin đã đăng nhập
Dữ liệu	Users, Posts, Comments, Flagged Posts, Trends, Emotion Distribution
Bộ lọc	Day, Week, Month, Year
Hậu điều kiện	-

2.3. Kiến trúc hệ thống

2.3.1. Kiến trúc tổng thể

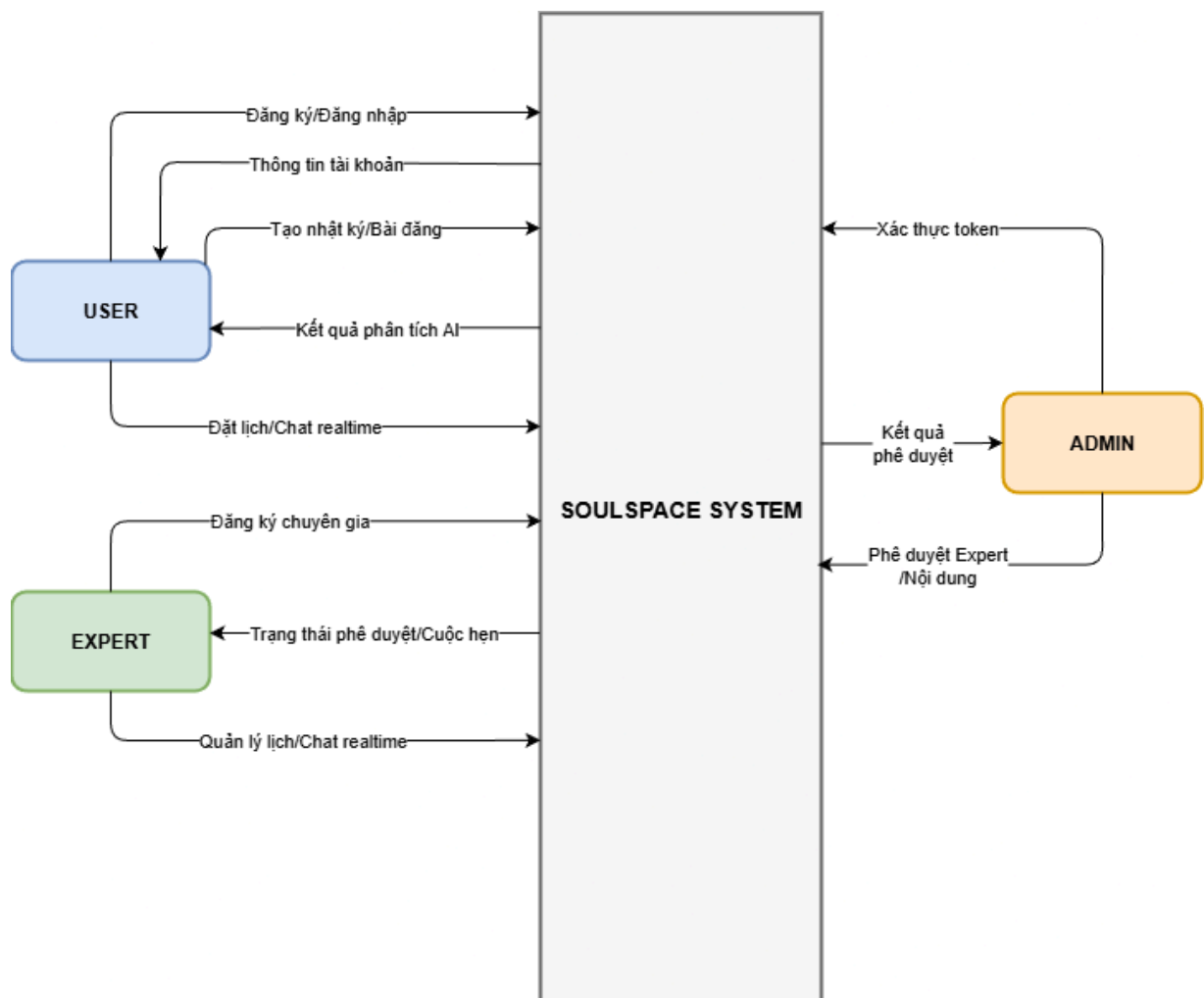


Hình 2.2: Kiến trúc tổng thể

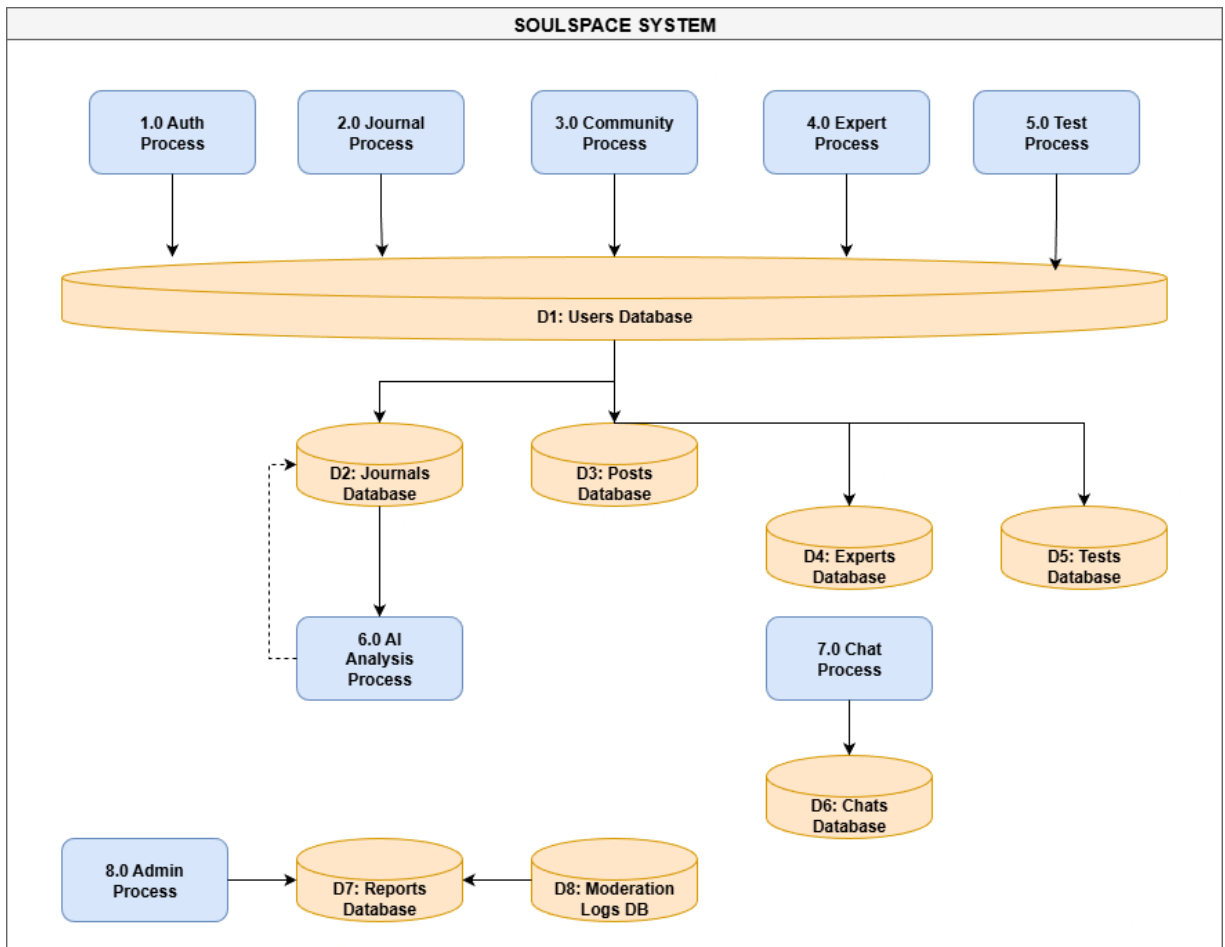
Đây là sơ đồ **Kiến trúc Dựa trên Dịch vụ (Service-Based Architecture)** được triển khai theo mô hình **Phân lớp (Layered Architecture)**.

- **Ưu điểm:** Phát triển cực nhanh ban đầu, dễ join dữ liệu, dễ quản lý transaction, kiến trúc đơn giản hơn Microservices rất nhiều. Phù hợp cho Startups/MVP.
- **Nhược điểm (Rủi ro):** Đây là một "nút cổ chai" (bottleneck). Nếu MongoDB bị quá tải, toàn bộ hệ thống đều sập. Việc thay đổi cấu trúc dữ liệu rất nhạy cảm vì có thể ảnh hưởng đến nhiều services cùng lúc
- **Khả năng mở rộng:** Tốt về mặt logic (code đã tách), nhưng sẽ gặp giới hạn về mặt dữ liệu (Database chung) khi scale lớn.

2.3.2. Data Flow Diagram (DFD)



Hình 2.3: Data Flow Diagram - Level 0 (Context Diagram)



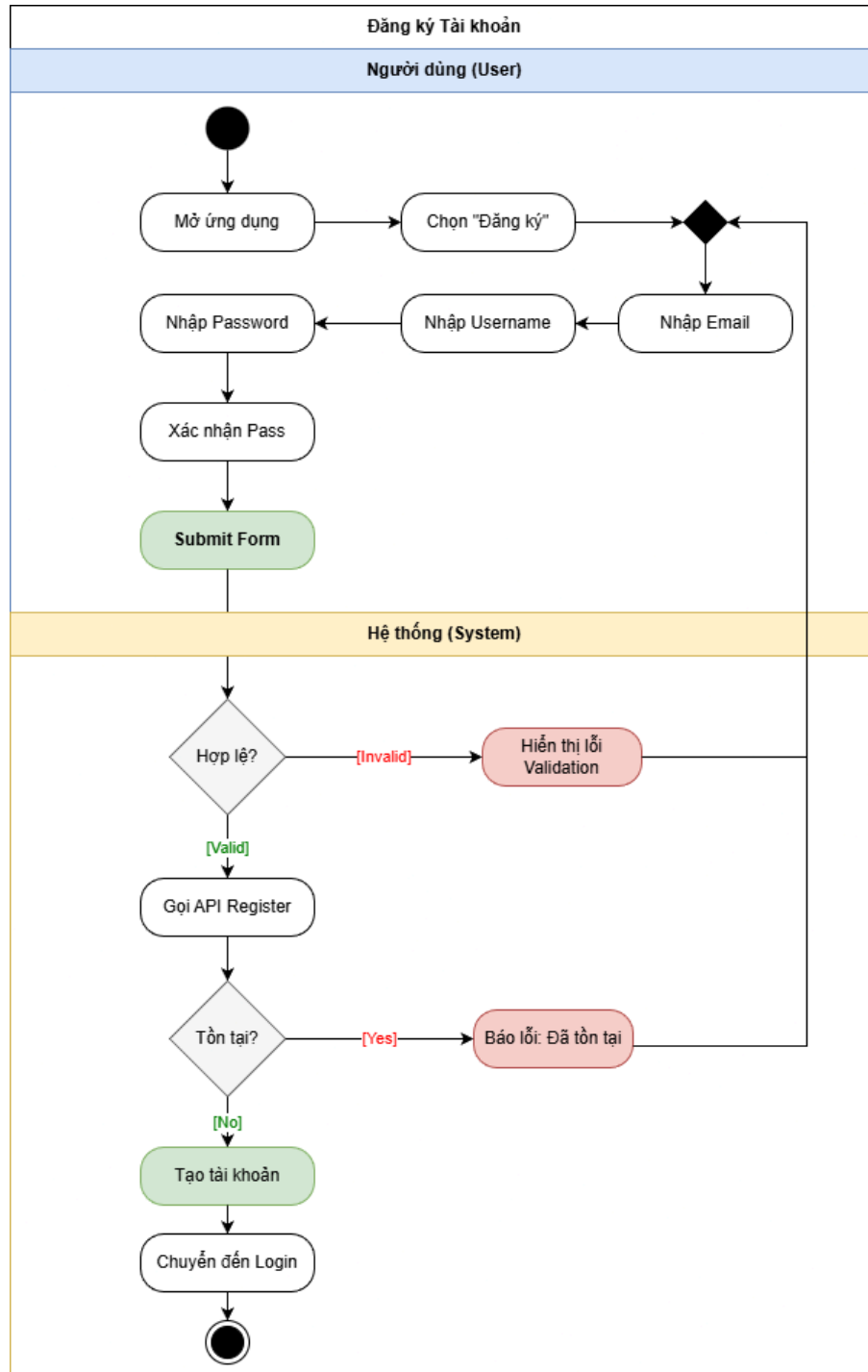
Hình 2.4: Data Flow Diagram - Level 1 (Context Diagram)

- **Bắt đầu (Định danh):** Người dùng hoặc Chuyên gia truy cập hệ thống qua **Quy trình Xác thực (1.0)**. Dữ liệu được lưu vào **Kho người dùng trung tâm (D1)**.
- **Chăm sóc bản thân (AI & Nhật ký):** Người dùng viết nhật ký qua **Quy trình Nhật ký (2.0)** lưu vào **Kho D2**. Ngay lập tức, **Quy trình AI (6.0)** sẽ đọc dữ liệu từ D2 để phân tích và trả về kết quả/lời khuyên tâm lý cho người dùng.
- **Tương tác xã hội & Chuyên gia:** Từ dữ liệu cá nhân, người dùng có thể tham gia cộng đồng (**Quy trình 3.0**) hoặc tìm kiếm chuyên gia (**Quy trình 4.0**).
- **Kết nối trực tiếp:** Nếu chọn chuyên gia, hệ thống kích hoạt **Quy trình Chat (7.0)** để hai bên nhắn tin, dữ liệu hội thoại được lưu tại **Kho D6**.

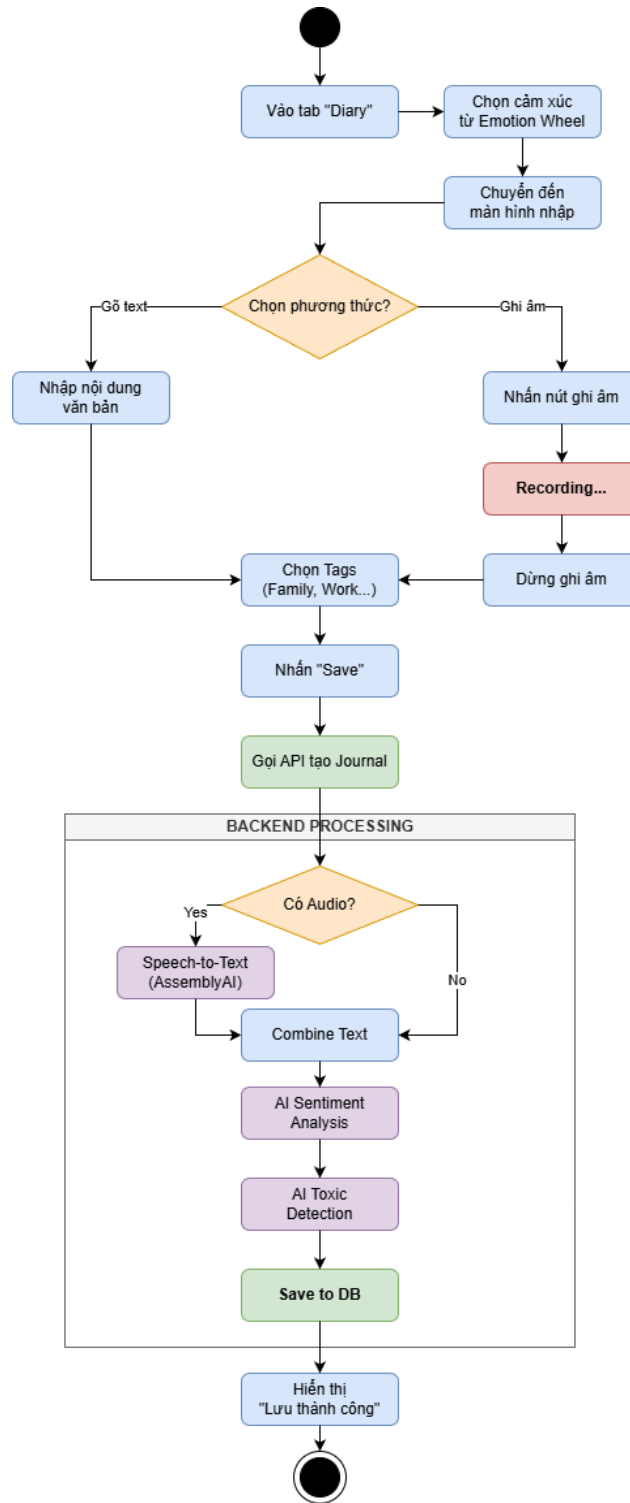
- **Giám sát & Quản lý:** Song song với các hoạt động trên, Admin thực hiện **Quy trình Quản trị (8.0)**, tiếp nhận báo cáo từ **Kho D7** và ghi nhận lịch sử kiểm duyệt vào **Kho D8** để đảm bảo an toàn cho hệ thống.

2.4. Activity Diagram

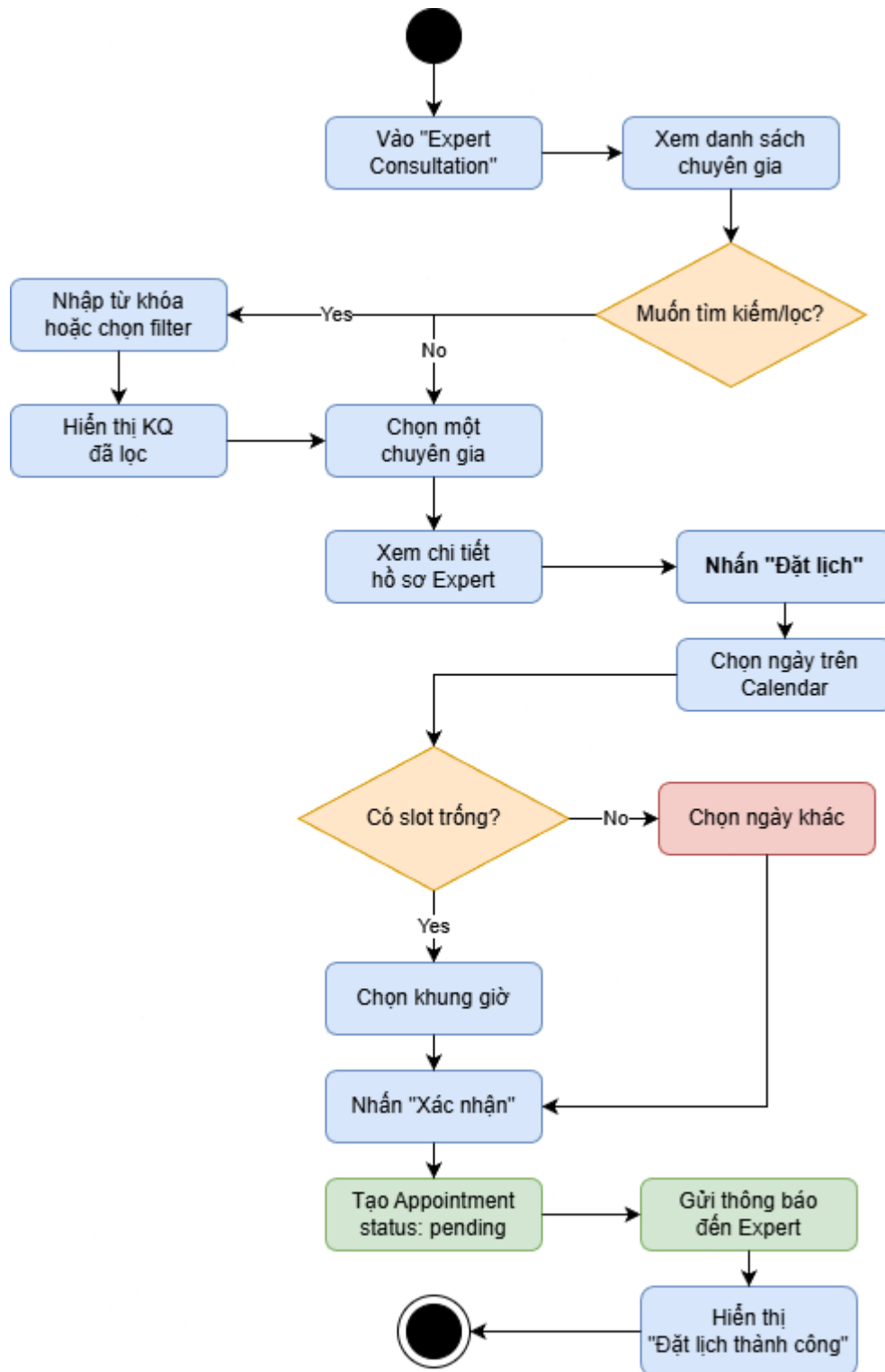
2.4.1. AD-01: Đăng ký Tài khoản User



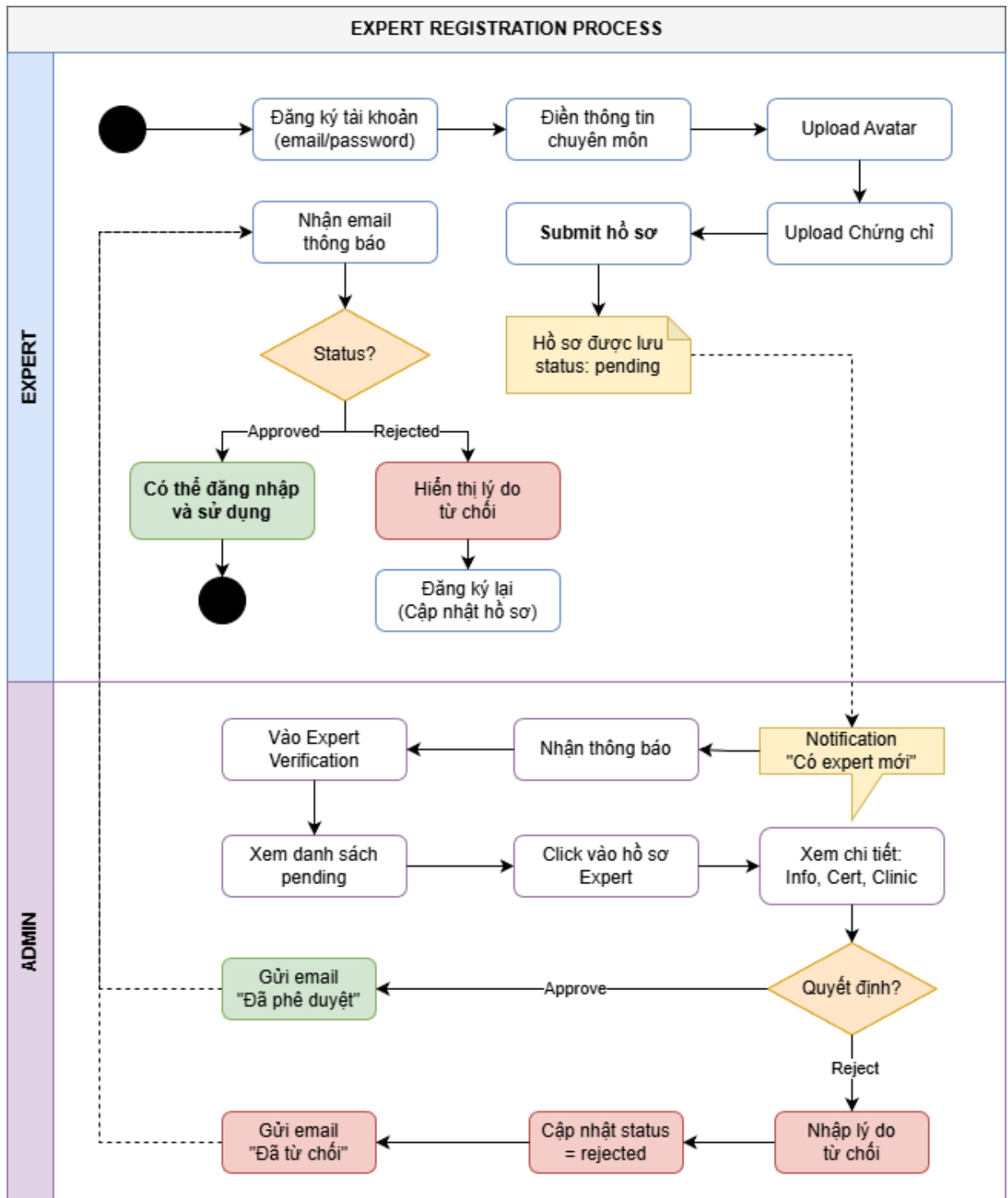
2.4.2. AD-02: Tạo Nhật ký Cảm xúc (Journal)



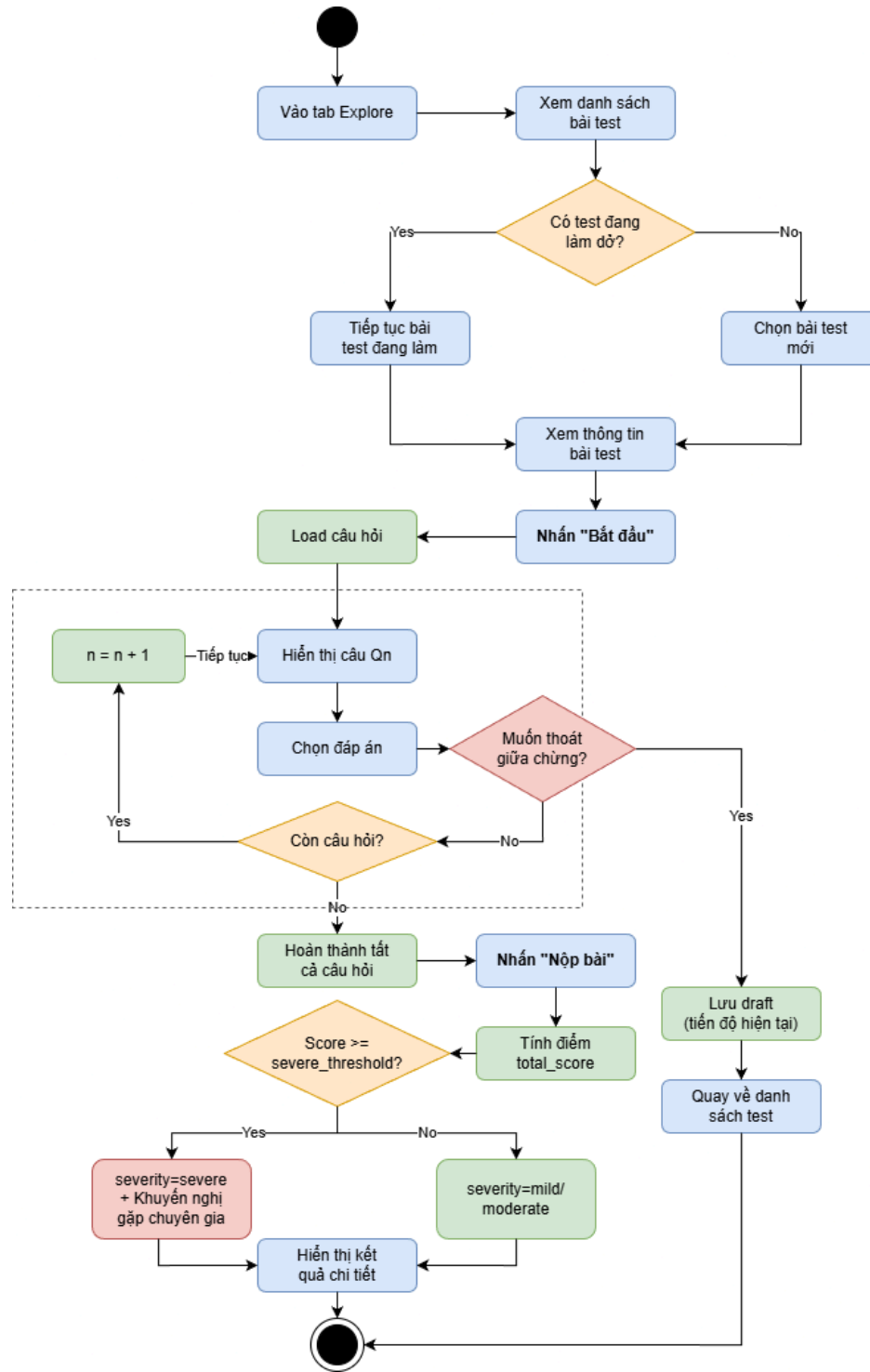
2.4.3. AD-03: Đặt lịch Tư vấn Chuyên gia



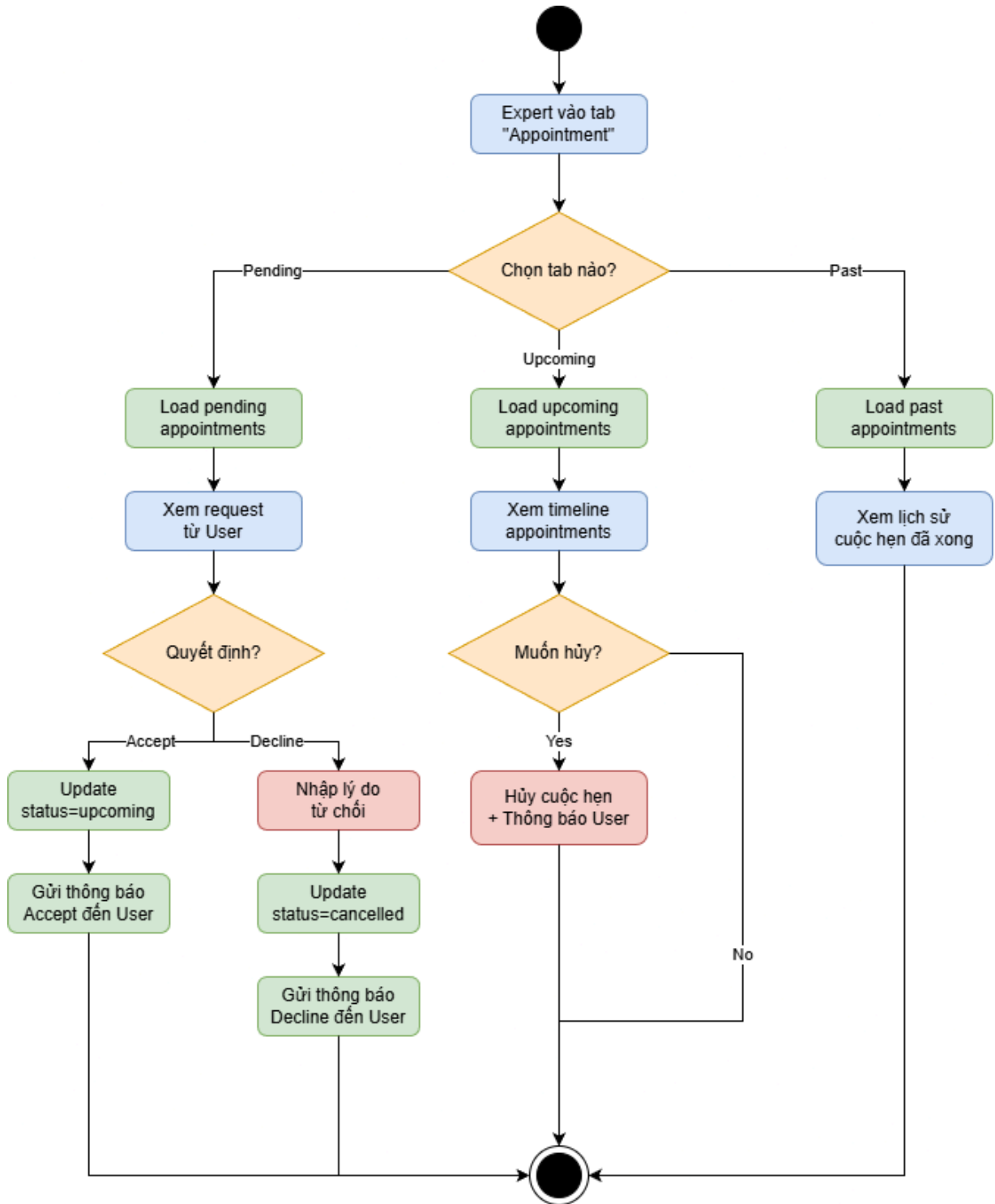
2.4.4. AD-04: Quy trình phê duyệt Expert



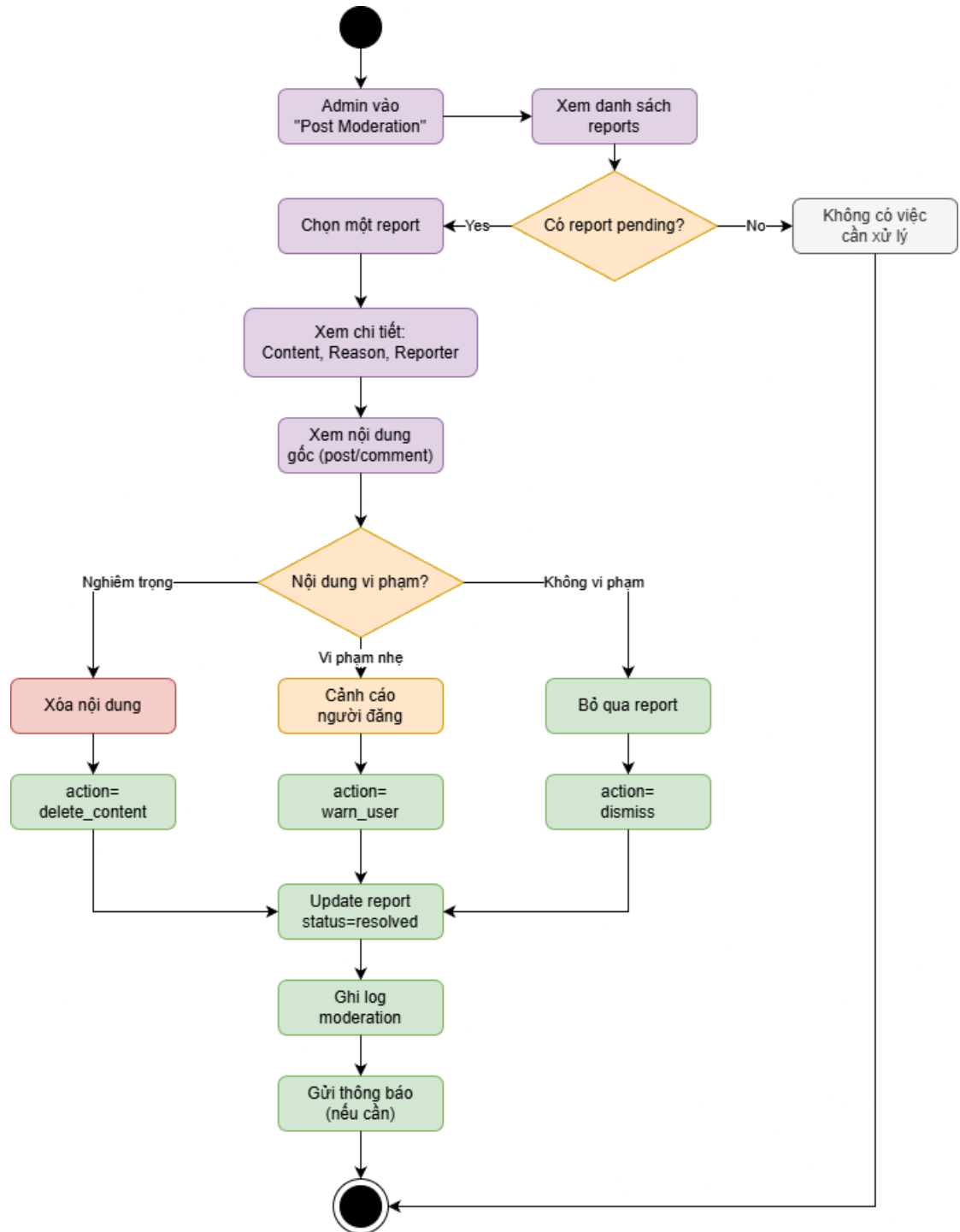
2.4.5. AD-05: Làm Bài Test Tâm lý



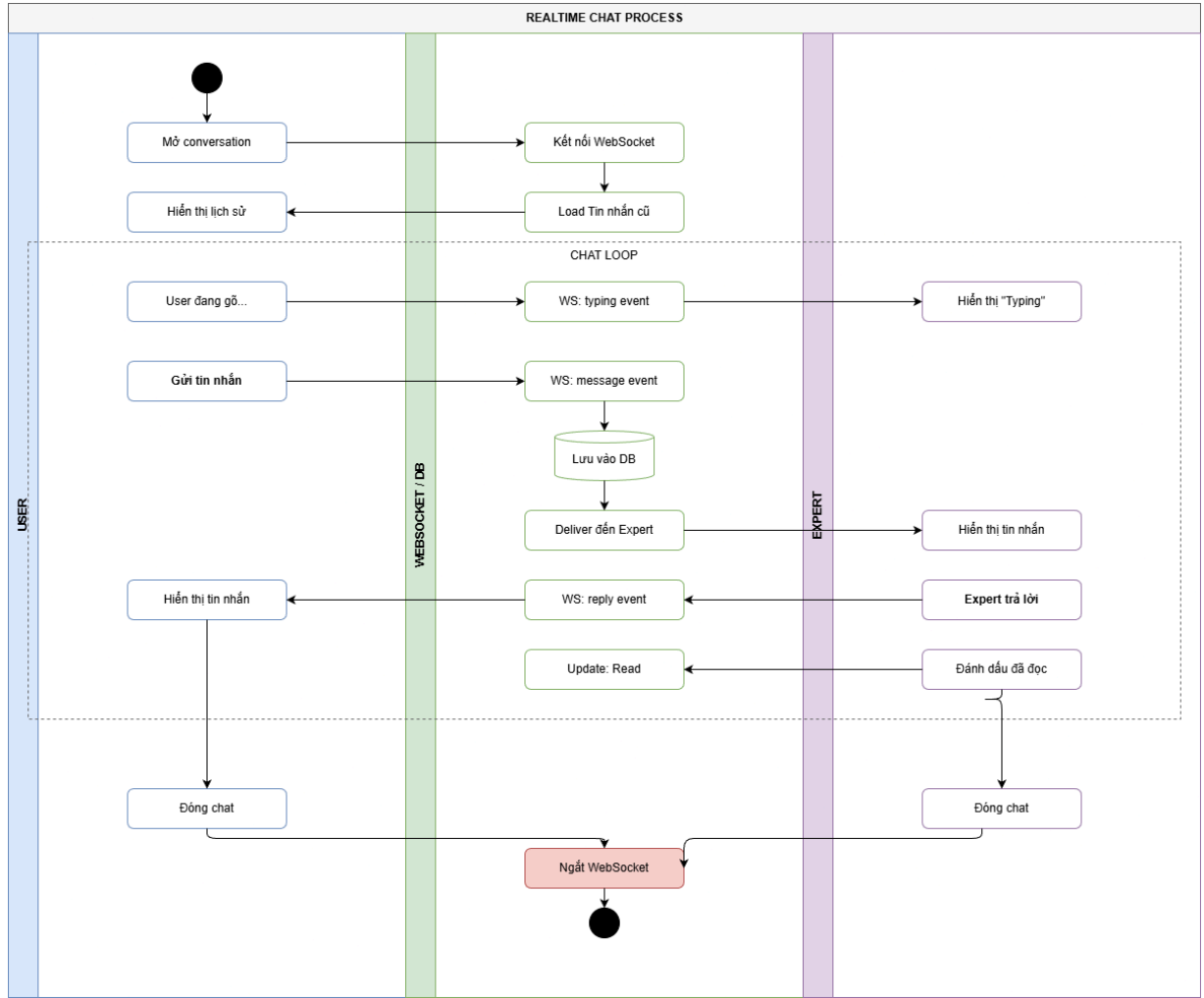
2.4.6. AD-06: Quản lý Cuộc hẹn (Expert)



2.4.7. AD-07: Kiểm duyệt Nội dung Cộng đồng (Admin)



2.4.8. AD-08: Chat Realtime giữa User và Expert



2.5. Một số giải thuật liên quan

2.5.1. Thuật toán nhận diện cảm xúc sử dụng mô hình Deep Learning (Tích hợp vào tính năng Journal)

Tổng quan: Sử dụng mô hình RoBERTa (Robustly Optimized BERT), được pre-trained 'roberta-base' (~125M params)

Pipeline xử lý: Thông qua 8 bước:

1. Load Data (original + augmented)
2. Preprocessing (normalize obfuscation, slang, context)
3. Train/Val Split (85%/15%)
4. Compute Class Weights
5. Tokenization (RoBERTa BPE)
6. Training (4 epochs, gradient accumulation)
7. Evaluation (F1, ROC-AUC)
8. Save Model

Giải thuật xử lý:

- **Label Mapping - Ánh xạ nhãn cảm xúc:** Model RoBERTa trả về các nhãn nội bộ (LABEL_0, LABEL_1, LABEL_2). Hệ thống ánh xạ sang nhãn có ý nghĩa: LABEL_0 → Negative, LABEL_1 → Neutral, LABEL_2 → Positive.

Mục đích sử dụng: Xử lý việc mất cân bằng của dữ liệu

- **Score Normalization - Chuẩn hóa điểm số:** Xác suất confidence (0-1) được chuyển thành score có dấu trong khoảng [-1.0, +1.0]:
 - Nếu sentiment = Positive: score = +confidence
 - Nếu sentiment = Negative: score = -confidence
 - Nếu sentiment = Neutral: score = 0.0

Mục đích sử dụng: Tạo thang đo trực quan - điểm dương = tích cực, điểm âm = tiêu cực, độ lớn = độ mạnh.

- **Confidence Classification - Phân loại độ tin cậy:** Dựa trên giá trị tuyệt đối của score để phân 3 mức:
 - $|\text{score}| \geq 0.8 \rightarrow$ "high" (Độ tin cậy cao)
 - $0.5 \leq |\text{score}| < 0.8 \rightarrow$ "medium" (Độ tin cậy trung bình)
 - $|\text{score}| < 0.5 \rightarrow$ "low" (Độ tin cậy thấp)

Mục đích sử dụng: Giúp hệ thống, người dùng biết khi nào tin tưởng kết quả.

- **Dual-Source Sentiment (cho Journal):** Journal hỗ trợ 2 nguồn phân tích song song:
 - **Icon-based:** Sử dụng ICON_SENTIMENT_MAP ánh xạ cứng từ icon người dùng chọn (Happy \rightarrow Positive/0.8, Sad \rightarrow Negative/-0.8, v.v.)
 - **AI-based:** Phân tích nội dung text thực tế qua model RoBERTa

Mục đích sử dụng: Icon-based cho kết quả nhanh, chính xác theo ý định chủ quan của người dùng. AI-based cho phân tích khách quan nội dung thực.

- **Phương pháp Lazy Loading - Tối ưu tài nguyên:** Model (~500MB) chỉ được load một lần duy nhất khi application startup. Sau đó, biến sentiment_pipeline global được tái sử dụng cho mọi request, tránh overhead load model nhiều lần, giúp giảm latency từ ~5s (load mỗi request) xuống ~100ms (inference only).
- **Sử dụng kỹ thuật Transfer Learning:** Sử dụng mô hình RoBERTa đã được pre-trained trên 160GB text data, sau đó fine-tune trên 58M tweets có gắn nhãn cảm xúc. Nhờ vậy, không cần train từ đầu mà vẫn đạt độ chính xác cao (~85% accuracy trên benchmark), có khả năng hiểu ngữ cảnh phức tạp như châm biếm, phủ định ("not good" \neq "good").

2.5.2. Thuật toán nhận diện văn bản Toxic (Tích hợp vào tính năng người dùng đăng bài)

Tổng quan: Sử dụng mô hình RoBERTa (Robustly Optimized BERT Approach), được fine-tune từ *cardiffnlp/twitter-roberta-base-sentiment* (~125M params). Đây là phiên

bản RoBERTa đã được huấn luyện đặc biệt trên ~58 triệu tweets để phân loại cảm xúc thành 3 nhãn: *Positive*, *Neutral*, *Negative*.

Ngoài ra, hệ thống còn hỗ trợ mô hình backup

distilbert-base-uncased-finetuned-sst-2-english (~66M params) - phiên bản nhẹ hơn được train trên tập Stanford Sentiment Treebank (SST-2).

Pipeline xử lý: Thông qua 8 bước:

1. Input Validation - Kiểm tra văn bản đầu vào (1-5000 ký tự, không rỗng)
2. Model Selection - Chọn model dựa trên config (RoBERTa hoặc DistilBERT)
3. Tokenization - Chuyển text thành token IDs sử dụng BPE (Byte-Pair Encoding)
4. Embedding - Chuyển tokens thành vectors 768 chiều + positional encoding
5. Transformer Encoding - Forward pass qua 12 encoder layers với self-attention
6. Classification - Project [CLS] token xuống 3 classes (Negative/Neutral/Positive)
7. Softmax - Chuyển logits thành phân phối xác suất
8. Post-processing - Label mapping + Score normalization (-1.0 đến +1.0)

Giải thuật xử lý:

- **Label Mapping - Ánh xạ nhãn cảm xúc:**

- Tính toán trọng số cho từng class dựa trên độ hiếm của nó: $weight = (total - positive) / positive$.
- Class càng hiếm, trọng số càng cao (được giới hạn trong khoảng [1.0, 3.0] để tránh lệch quá đà).
- Khi tính loss, lỗi sai ở class hiếm sẽ xảy ra nhiều hơn, buộc model phải học kỹ các trường hợp này.

Mục đích sử dụng: Xử lý việc mất cân bằng của dữ liệu

- **Tối ưu hóa Hiệu năng (Performance Optimization):** Để train được model bert-base trên phần cứng hạn chế (VRAM thấp), nhóm đã sử dụng phương pháp:

- Mixed Precision (FP16): Tính toán ở độ chính xác 16-bit giúp giảm 50% bộ nhớ VRAM.
- Gradient Accumulation: Thay vì cập nhật trọng số sau mỗi batch nhỏ (8 samples), hệ thống tích lũy gradient qua 2 bước rồi mới cập nhật, mô phỏng việc training với batch size lớn hơn (16 samples) để ổn định việc học.
- **Sử dụng kỹ thuật Transfer Learning:** Sử dụng mô hình RoBERTa có sẵn, tinh chỉnh mô hình lại để tạo ra mô hình có độ chính xác cao, nhận diện được lên đến 80% bài viết.

2.5.3. Thuật toán liên quan đến hệ thống Cây Tinh Thần (User Tree System)

Tổng quan: Hệ thống User Tree là một tính năng cốt lõi của SoulSpace, nơi các hành động chăm sóc bản thân (Positive Actions) của người dùng sẽ được chuyển hóa thành "dinh dưỡng" để nuôi lớn một cái cây ảo. Sử dụng cơ chế XP (Kinh nghiệm) lũy tiến và Streak (Chuỗi ngày) để khuyến khích thói quen tích cực hàng ngày.

- **Mục tiêu:** Khuyến khích người dùng quay lại ứng dụng mỗi ngày.
- **Cơ chế:** Mỗi ngày người dùng thực hiện hành động tích cực sẽ nhận được XP. Cây sẽ lớn lên (tăng Level) khi đủ XP. Streak sẽ tăng nếu người dùng thực hiện tiêu tiếp.

Pipeline xử lý: Thông qua 6 bước:

1. Validation Check - Kiểm tra hành động (Action ID) có tồn tại trong hệ thống.
2. Daily Limit Check - Kiểm tra giới hạn: Người dùng đã tưới cây hôm nay chưa? (Chỉ cho phép 1 lần/ngày).
3. XP Calculation - Cộng thêm lượng XP cố định ($XP_PER_ACTION = 10$) vào tổng XP của cây.
4. Streak Calculation - Tính toán chuỗi ngày liên tiếp dựa trên ngày tưới gần nhất và ngày hiện tại.

5. Level Calculation - Từ tổng XP mới, tính toán Level hiện tại và tiến độ đến Level tiếp theo dựa trên thuật toán lũy tiến.

6. Persistence & Response - Lưu dữ liệu mới xuống database (UserTree & User Points) và trả về trạng thái cây mới nhất cho client.

Giải thuật xử lý:

- **Daily Throttling - Giới hạn tần suất:** Kiểm tra tính hợp lệ của hành động dựa trên thời gian UT bằng cách lấy ngày hiện tại (`today_utc`) và ngày tưới lần cuối (`last_watered_date`).

- Nếu `last_watered_date` trùng với `today_utc` => Chặn hành động (Return Error: "Already watered today").
- Nếu `last_watered_date` < `today_utc` hoặc chưa từng tưới => Cho phép.

Mục đích sử dụng: Ngăn chặn việc spam hành động để cày điểm, buộc người dùng phải quay lại vào ngày hôm sau, từ đó hình thành thói quen lâu dài.

- **Streak Tracking - Tính chuỗi ngày liên tiếp:** So sánh ngày hiện tại với ngày tưới lần cuối để xác định chuỗi streak:
 - Nếu tưới vào ngày hôm qua (Yesterday): $\text{New Streak} = \text{Current Streak} + 1$ (Duy trì chuỗi).
 - Nếu tưới vào ngày hôm nay (Today): $\text{New Streak} = \text{Current Streak}$ (Không đổi).
 - Nếu tưới vào ngày kia trở về trước ($\text{Gap} > 1$ day): $\text{New Streak} = 1$ (Reset chuỗi, bắt đầu lại).

Mục đích sử dụng: Mục đích sử dụng: Tạo động lực tâm lý "sợ mất chuỗi" (Loss Aversion), khuyến khích người dùng duy trì mở ứng dụng hàng ngày.

- **Progressive Leveling - Thuật toán lũy tiến:** Sử dụng cơ chế "càng lên cao càng khó" (Progressive Difficulty) để tính Level.
 - **Base XP:** Level 1 lên Level 2 cần 50 XP.
 - **Scaling:** Mỗi level tiếp theo cần thêm 25 XP so với level trước đó.

- Lv 1 → Lv 2: Cần 50 XP
- Lv 2 → Lv 3: Cần $50 + 25 = 75$ XP
- Lv 3 → Lv 4: Cần $75 + 25 = 100$ XP
- **Giải thuật:** Sử dụng vòng lặp (while loop) để trừ dần tổng XP tích lũy (total_xp) vào các mốc XP cần thiết cho đến khi số dư không đủ để lên cấp tiếp theo. Phần dư chính là current_xp_in_level.

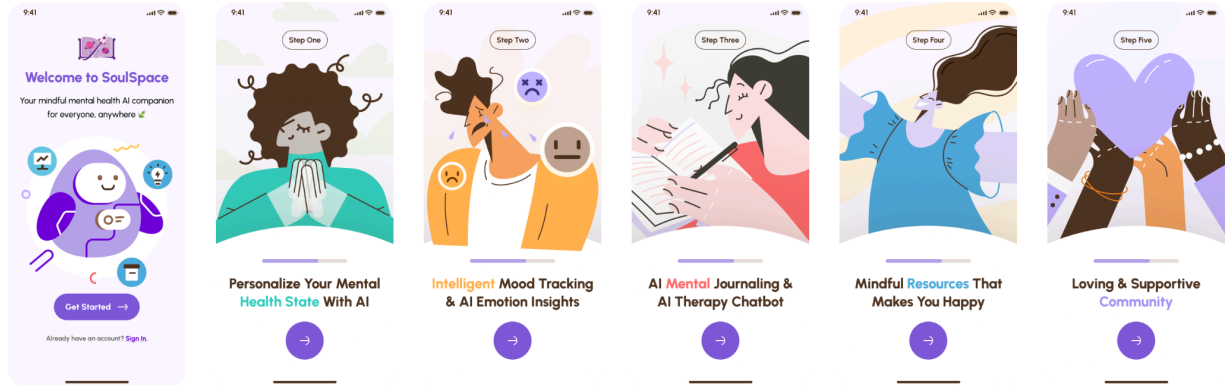
Mục đích sử dụng: Tạo cảm giác tiến bộ nhanh ở giai đoạn đầu (người dùng mới) để giữ chân họ, và tạo thử thách dài hạn ở giai đoạn sau (người dùng lâu năm) để duy trì sự hứng thú.

- **Data Synchronization:** Khi cây được tưới, hệ thống cập nhật đồng thời 2 nơi dữ liệu:
 - **UserTree Collection:** Lưu trạng thái cây (total_xp, level, streak, actions_history).
 - **User Collection:** Cộng dồn vào điểm thưởng chung (total_points) của người dùng (dùng để đổi quà hoặc ranking).

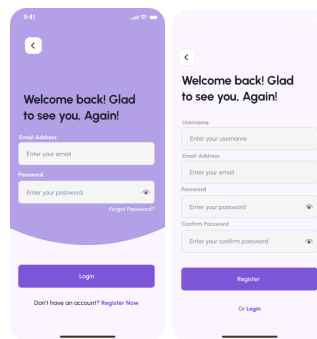
Mục đích sử dụng: Đảm bảo nhất quán dữ liệu giữa minigame và hệ thống profile chính của người dùng.

2.6. Danh sách các màn hình phác thảo dự kiến

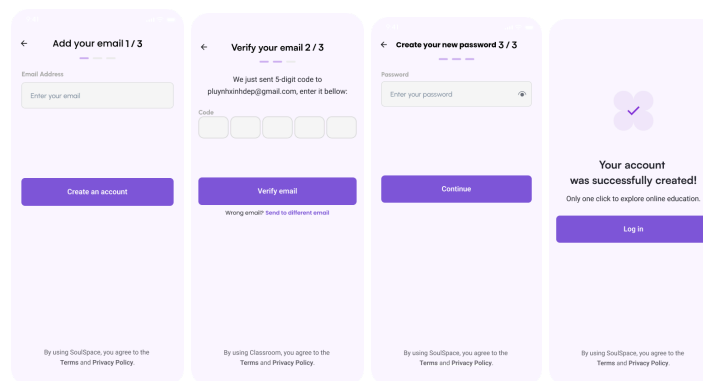
2.6.1. Nhóm màn hình liên quan đến Xác thực (Authentication)



Hình 2.6.1: Màn hình chào mừng



Hình 2.6.2: Màn hình đăng ký, đăng nhập

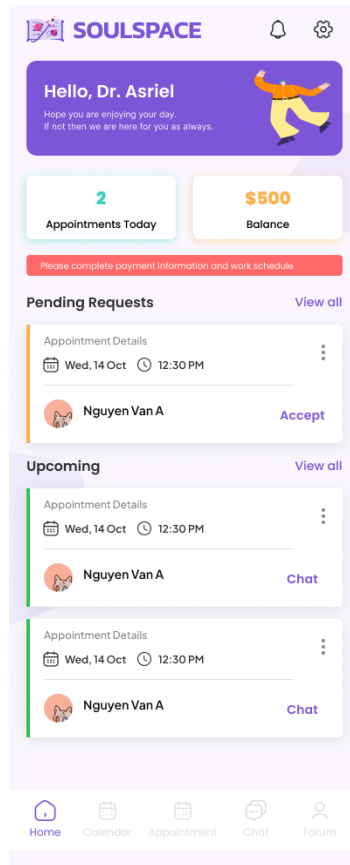


Hình 2.6.3: Màn hình quên mật khẩu

2.6.2. Nhóm màn hình liên quan đến màn hình chính

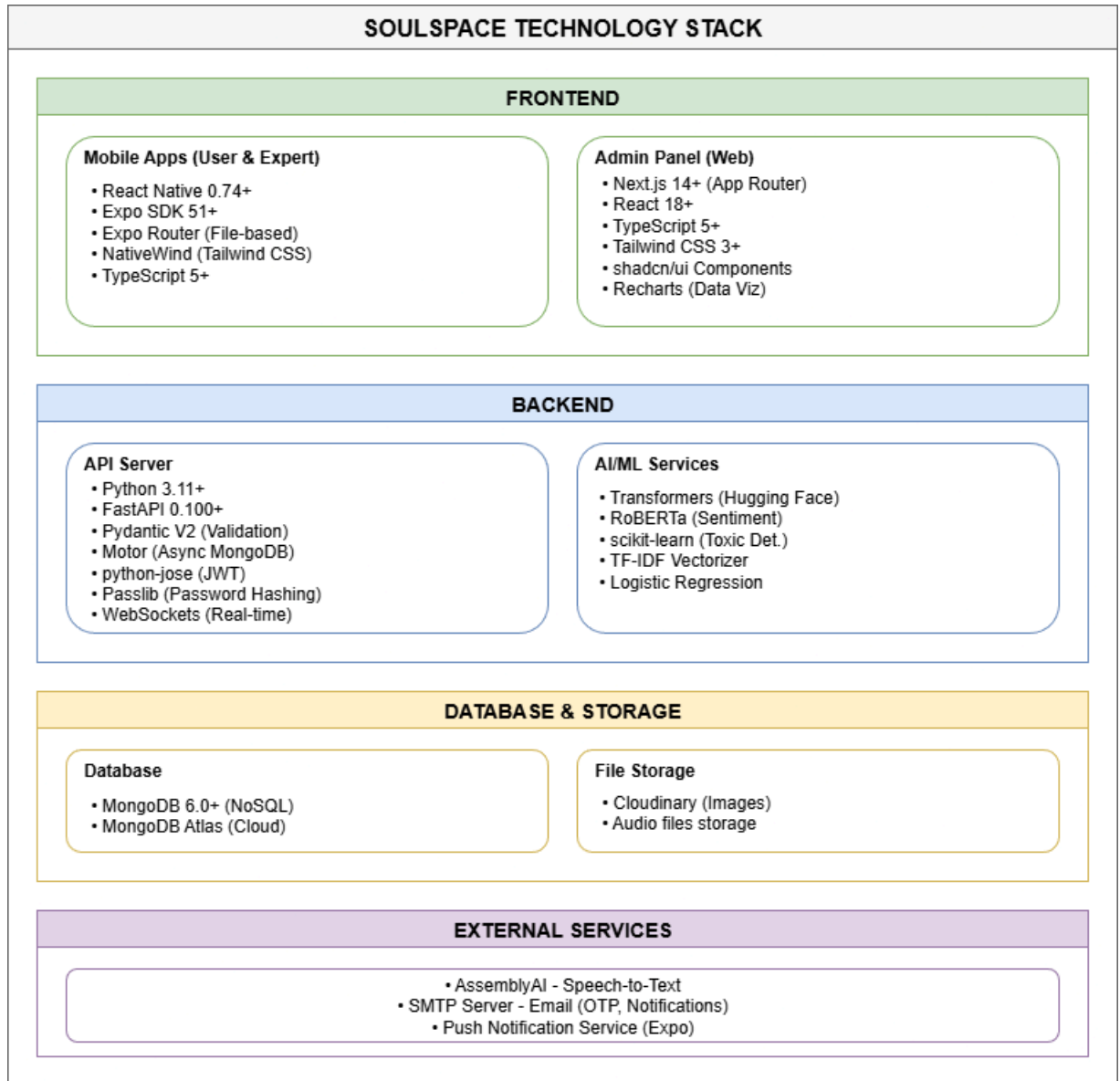


Hình 2.6.4: Màn hình giao diện chính (dành cho User)



Hình 2.6.5: Màn hình giao diện chính (dành cho Expert)

2.7. Công nghệ sử dụng



Hình 2.: Công nghệ sử dụng

2.7.1. Backend (SoulSpace-Backend)

Core Framework & Runtime

Công nghệ	Phiên bản	Mục đích	Lý do lựa chọn
Python	3.11+	Runtime	Hiệu năng cao, typing support, ecosystem ML/AI phong phú
FastAPI	0.100+	Web Framework	Async native, auto OpenAPI docs, type hints, hiệu năng cao
Uvicorn	0.23+	ASGI Server	Async server hiệu năng cao cho FastAPI
Pydantic	2.0+	Data Validation	Validation mạnh mẽ, serialization, settings management

Database & ORM

Công nghệ	Phiên bản	Mục đích	Lý do lựa chọn
MongoDB	6.0+	Database	Schema flexible, scale horizontal, phù hợp cho data đa dạng
Motor	3.3+	Async Driver	Driver MongoDB async native cho Python
Beanie (optional)	1.23+	ODM	Object Document Mapper cho MongoDB

Authentication & Security

Công nghệ	Phiên bản	Mục đích	Lý do lựa chọn
python-jose	3.3+	JWT	Tạo và verify JSON Web Tokens
passlib	1.7+	Password	Hashing mật khẩu an toàn (bcrypt)
python-multipart	0.0.6+	File Upload	Xử lý multipart form data

AI/ML Libraries

Công nghệ	Phiên bản	Mục đích	Lý do lựa chọn
transformers	4.35+	NLP Models	Hugging Face transformers cho sentiment analysis
torch	2.0+	Deep Learning	Backend cho transformers models
scikit-learn	1.3+	ML Classic	TF-IDF, Logistic Regression cho toxic detection
joblib	1.3+	Model Serialization	Lưu/load ML models

External Services Integration

Công nghệ	Phiên bản	Mục đích	Lý do lựa chọn
cloudinary	1.36+	Image Storage	CDN, image optimization, transformation

Công nghệ	Phiên bản	Mục đích	Lý do lựa chọn
assemblyai	0.20+	Speech-to-Text	API dễ dùng, hỗ trợ tiếng Việt
aiosmtp	2.0+	Email	Async email sending
httpx	0.25+	HTTP Client	Async HTTP client

Utilities

Công nghệ	Phiên bản	Mục đích	Lý do lựa chọn
python-dotenv	1.0+	Environment	Load environment variables
APScheduler	3.10+	Job Scheduling	Background tasks scheduling

2.7.2. Mobile Apps (SoulSpace & SoulSpace-FE-Expert)

Core Framework

Công nghệ	Phiên bản	Mục đích	Lý do lựa chọn
React Native	0.74+	Mobile Framework	Cross-platform (iOS/Android), large ecosystem
Expo	SDK 51+	Development Platform	Faster development, OTA updates, managed workflow
TypeScript	5.0+	Language	Type safety, better developer experience

Navigation & Routing

Công nghệ	Phiên bản	Mục đích	Lý do lựa chọn
Expo Router	3.0+	Navigation	File-based routing, giống Next.js, type-safe
React Navigation	6.0+	Navigation Core	Underlying navigation library

Styling & UI

Công nghệ	Phiên bản	Mục đích	Lý do lựa chọn
NativeWind	4.0+	Styling	Tailwind CSS cho React Native, utility-first
Tailwind CSS	3.4+	CSS Framework	Rapid UI development, consistent design
Lucide React Native	0.300+	Icons	Modern icon library
React Native Reanimated	3.6+	Animations	Smooth, performant animations

State Management & Data

Công nghệ	Phiên bản	Mục đích	Lý do lựa chọn
React Query/TanStack	5.0+	Data Fetching	Caching, background sync, optimistic updates

Công nghệ	Phiên bản	Mục đích	Lý do lựa chọn
Zustand (optional)	4.4+	State Management	Simple, lightweight state management
AsyncStorage	1.21+	Local Storage	Persistent local data
Expo SecureStore	13.0+	Secure Storage	Secure token storage

Media & Device

Công nghệ	Phiên bản	Mục đích	Lý do lựa chọn
Expo Image Picker	15.0+	Image Selection	Camera/gallery access
Expo AV	14.0+	Audio/Video	Recording, playback
Expo Notifications	0.25+	Push Notifications	Local & remote notifications
Expo Haptics	13.0+	Haptic Feedback	Vibration feedback

Utilities

Công nghệ	Phiên bản	Mục đích	Lý do lựa chọn
date-fns	3.0+	Date Handling	Date manipulation utilities
Formik/React Hook Form	-	Form Management	Form validation & handling

Công nghệ	Phiên bản	Mục đích	Lý do lựa chọn
Yup/Zod	-	Validation	Schema validation

2.7.3. Admin Panel (SoulSpace-FE-Admin)

Core Framework

Công nghệ	Phiên bản	Mục đích	Lý do lựa chọn
Next.js	14+	React Framework	App Router, SSR/SSG, API routes, optimized
React	18+	UI Library	Component-based, hooks, concurrent features
TypeScript	5.0+	Language	Type safety, better maintainability

UI Components & Styling

Công nghệ	Phiên bản	Mục đích	Lý do lựa chọn
shadcn/ui	-	UI Components	High-quality, customizable, accessible
Radix UI	-	Primitives	Headless, accessible UI primitives
Tailwind CSS	3.4+	Styling	Utility-first, rapid development

Công nghệ	Phiên bản	Mục đích	Lý do lựa chọn
Lucide React	0.300+	Icons	Modern, consistent icons
class-variance-authority	0.7+	Variant Management	Component variants
clsx / tailwind-merge	-	Class Utilities	Conditional classes

Data Visualization

Công nghệ	Phiên bản	Mục đích	Lý do lựa chọn
Recharts	2.10+	Charts	React-based charting library
@tanstack/react-table	8.10+	Data Tables	Powerful, flexible tables

Form & Validation

Công nghệ	Phiên bản	Mục đích	Lý do lựa chọn
React Hook Form	7.48+	Form Management	Performant forms
Zod	3.22+	Validation	TypeScript-first validation
@hookform/resolvers	3.3+	Resolver	Zod integration with RHF

Utilities

Công nghệ	Phiên bản	Mục đích	Lý do lựa chọn
date-fns	3.0+	Date Handling	Date manipulation
sonner	1.2+	Toast	Toast notifications
next-themes	0.2+	Theme	Dark/Light mode

2.7.4. External Services & APIs

Service	Provider	Mục đích
MongoDB Atlas	MongoDB	Database hosting
Cloudinary	Cloudinary	Image/media storage & CDN
AssemblyAI	AssemblyAI	Speech-to-Text API
Expo EAS	Expo	Build & submit apps

Chương 3. ĐÁNH GIÁ VÀ KẾT LUẬN

3.1. Đánh giá mức độ hoàn thành

Tính đến thời điểm viết báo cáo, nhóm thực hiện đã hoàn thành khoảng 85% khối lượng công việc so với đề cương chi tiết ban đầu, bao gồm:

- Đã hoàn thiện toàn bộ Backend API cho 3 role (User, Expert, Admin).
- Tích hợp AI: Phân tích cảm xúc, Speech-to-Text, Lọc nội dung độc hại.
- Cơ sở dữ liệu MongoDB được thiết kế tối ưu cho truy vấn phức tạp.

3.2. Phương hướng trong tương lai

Khắc phục các tính năng chưa đạt được, và phát triển thêm một số tính năng mới:

- **Tích hợp wearable (Apple Watch, Mi Band...)** để theo dõi giấc ngủ, nhịp tim và dự đoán cảm xúc chính xác hơn.
- **Cộng đồng peer support riêng theo chủ đề** (ví dụ: sinh viên năm cuối, người mất phương hướng nghề nghiệp, LGBTQ...)
- **Tủ sách & podcast chữa lành**, phân loại theo vấn đề (stress học tập, lo âu xã hội...)
- **Hợp tác với trường học & doanh nghiệp** để triển khai phiên bản chăm sóc tinh thần cho học sinh/sinh viên và nhân viên.

3.3. Kết luận

- User: Có người bạn đồng hành 24/7, không gian an toàn để chữa lành.
- Expert: Nền tảng tiếp cận bệnh nhân trẻ, quản lý công việc hiệu quả.
- Xã hội: Góp phần nâng cao nhận thức và hỗ trợ sức khỏe tinh thần cho thế hệ Gen Z.

TÀI LIỆU THAM KHẢO